

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-074129

(43)Date of publication of application : 17.03.1998

(51)Int.Cl.

G06F 3/06

G06F 3/06

G06F 3/06

G06F 3/06

(21)Application number : 08-229692

(71)Applicant : HITACHI LTD

(22)Date of filing : 30.08.1996

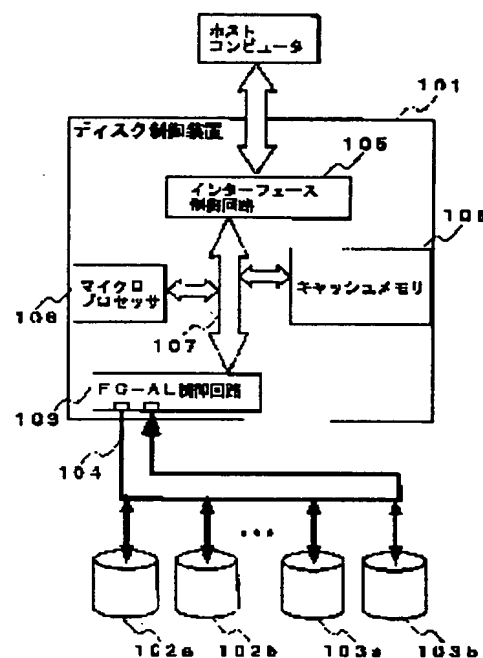
(72)Inventor : FUJIBAYASHI AKIRA

(54) DISK ARRAY DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To reduce the quantity of data which are transferred between a data duplexing disk controller and a magnetic disk device by writing the data into a 1st disk device from the disk controller via an interface and transferring the data to a 2nd disk device of a same redundant constitution group existing on the interface via the 1st disk device.

SOLUTION: The write data are written in a cache memory 108, and a microprocessor 106 sends the data of the memory 108 to an FC-AL control circuit 109 based on the write request parameter of a host computer. Then it is instructed to write the data given to the circuit 109 into a magnetic disk device 102a. At the same time, the circuit 109 produces a write command and issues it via an FC-AL interface. Thus, the data are sent to the device 102a from the host computer. The device 102a writes the received data and also transfers them to a magnetic disk device 102b which is paired with the device 102a.



LEGAL STATUS

[Date of request for examination]

04.10.2000

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Best Available Copy

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平10-74129

(43)公開日 平成10年(1998)3月17日

(51)Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 3/06	5 4 0		G 0 6 F 3/06	5 4 0
	3 0 1			3 0 1 X
	3 0 2			3 0 2 E
	3 0 4			3 0 4 F

審査請求 未請求 請求項の数 2 O L (全 9 頁)

(21)出願番号 特願平8-229692

(22)出願日 平成8年(1996)8月30日

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72)発明者 藤林 昭

東京都国分寺市東恋ヶ窪一丁目280番地

株式会社日立製作所中央研究所内

(74)代理人 弁理士 小川 勝男

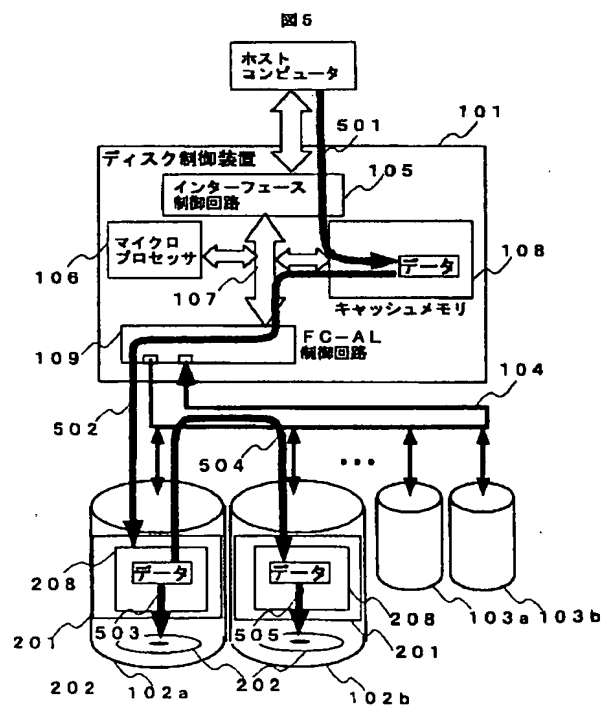
(54)【発明の名称】 ディスクアレイ装置

(57)【要約】

【課題】 RAID 1 の冗長構成を用いるディスクアレイ装置において、データ書き込みの際、データを二重化する必要から発生するデータの転送量増加を抑える。

【解決手段】 RAID 1 の冗長構成を用いるディスクアレイ装置において、複数の磁気ディスク装置とディスク制御装置を接続するインターフェースを設け、データ書き込み際に、第一の磁気ディスク装置にデータ転送し、該データを第一の磁気ディスク装置が、ディスク制御装置を経由せずに、第二の磁気ディスク装置に直接転送する。

【効果】 データ書き込みの際、ディスク制御装置と磁気ディスク装置間のデータ転送量が減少して、ディスク制御装置の処理性能が向上する。



【特許請求の範囲】

【請求項1】複数のディスク装置と、ホストコンピュータからの入出力要求をうけて該ディスク装置に対するデータ入出力を制御するディスク制御装置と、該複数のディスク装置と該ディスク制御装置を接続するインターフェースからなり、該ディスク装置の内の複数台が一組となりデータを重複して格納する冗長構成を用いるディスクアレイ装置であって、

データの書き込み時に、該ディスク制御装置から第一のディスク装置に該インターフェースを介してデータを書き込む手段と、

該第一のディスク装置が、該第一のディスク装置が接続しているインターフェース上に存在する同一の冗長構成グループの第二のディスク装置に該データを転送する手段とを有するディスクアレイ装置。

【請求項2】該ディスク装置は、

複数のコマンド及びデータを一時的に保持して置く一時記憶手段と、

他のコマンド実行中に、同一の冗長構成グループのディスク装置からのコマンドおよびデータを該一時記憶手段にて受信および保持する手段と、

該一時記憶手段に保持されている実行待ちの複数のコマンドの実行順序を変更する手段を備える請求項1記載のディスクアレイ装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、ディスクアレイ装置に関し、特に、データに二重化時のディスクアレイ装置内のディスク制御装置とディスク装置との間のインターフェース上のデータ転送量を低減することで、ディスク制御装置の処理性能を向上させるディスクアレイ装置に関する。

【0002】

【従来の技術】記憶装置の信頼性の向上を目的としたディスクアレイ装置では、データの読出し、書き込みが、故障により不可能になるのを防ぐため、データを複数の磁気ディスク装置に多重化して記録する冗長構成を用いている。数種の冗長構成が、「ACase for Redundant Array of Inexpensive Disk(RAID)」(Proceedings of the 1988 ACM SIGMOD Conference on the Management of Data 1988)に記載されている。この論文で規定されている数種の冗長構成の中で、とくに二台の磁気ディスク装置にデータを重複して書き込む、RAID 1の呼称で規定されている冗長構成の場合、二台の内の一方の磁気ディスク装置が故障しても、もう一方の正常な磁気ディスク装置に同一のデータが記録されているので、データの読出しは可能である。

【0003】図8は、従来のデータ二重化の方式を用いるディスクアレイ装置の構成の一例を示した概略図である。ディスクアレイ装置は、ディスク制御装置801、

及びディスク装置802a、802b、803a、803bから構成される。ディスク制御装置801は、ディスク装置との通信路を制御するSCSI制御回路809a、809b、809c、809d、ホストコンピュータとのインターフェースであるインターフェース制御回路805、マイクロプロセッサ806、キャッシュメモリ808から構成される。また、図9はこのような従来のディスクアレイ装置における新データ書き込みのシーケンスを示すフローチャートである。図8および図9を用いて、従来のデータ二重化の方式を用いるディスクアレイ装置における、ホストコンピュータからデータ書き込み、読み出しの要求に対する動作について以下に説明する。

【0004】まず、書き込み要求の場合は、ホストコンピュータからの書き込み要求を受けたディスクアレイ装置内のディスク制御装置801が、二台一組で冗長構成となっている磁気ディスク装置802a、802bの両方に対してWRITEコマンドまたはパラメータを送り(ステップ901)、該WRITEコマンドまたはパラメータを受けて、それぞれの磁気ディスク装置802a、802bは、磁気ヘッドの位置をディスク上のデータを転送すべき位置に移動させ、またデータを書き込むべき位置がヘッドの直下に回転してくるまで待つシーク・サーチ動作を開始する(ステップ902)。磁気ディスク装置802a、802bが、このシーク・サーチ動作を行っているのと平行して、ディスク制御装置801は該磁気ディスク装置内のバッファメモリに該データを転送してデータ転送処理を終了する(ステップ903)。シーク・サーチ動作が完了すると、該磁気ディスク装置802a、802bは、バッファメモリにある該データをディスクに書き込み(ステップ904)、書き込みが正常に終了したことをディスク制御装置801に通知する。この通知を受けたディスク制御装置801は二台の磁気ディスク装置がそれぞれ正常に書き込みを行ったことを確認し、ホストコンピュータに書き込み要求が正常に終了したことを通知する。

【0005】次に、ホストコンピュータからデータの読み出し要求を受けた場合であるが、この場合には二通りの方法がある。

【0006】第一の方法は、ディスクアレイ装置内のディスク制御装置801は、あらかじめ決めてある一方の磁気ディスク装置802aに、読み出しコマンドおよびパラメータを発行する。これを受けた磁気ディスク装置802aは、シーク・サーチ動作を開始し、該データの転送が可能な状態になると磁気ディスク装置802a内のバッファメモリにデータを送る。また、磁気ディスク装置802aは、バッファメモリに該データを転送する動作と平行して、データ転送可能状態であることをディスク制御装置801に通知し、この通知を受けたディスク制御装置801は、この磁気ディスク装置802a内のバッファメモリから該データを読み出し、ホストコン

ピュータにデータ転送を行う。

【0007】第二の方法は、冗長構成となっている二台一組の磁気ディスク装置のシーク・サーチ動作の早く完了する方にコマンドを出す方法である。この方法の場合、二台の磁気ディスク装置の早くデータ転送可能状態になる方の磁気ディスク装置を選ぶので、磁気ディスク装置の読み出し時間を短くすることが出来る。ただし、この方法の場合は、ディスク制御装置が、それぞれの磁気ディスク装置にすでに発行されているコマンドの待ち行列を管理している必要がある。

【0008】ここで、従来のデータ二重化の方式における問題として挙げられることは、新データ書き込みの処理の場合、同一データを二台の磁気ディスクに書き込まなければならないので、単純に通常の磁気ディスク装置に書き込む場合の2倍のマイクロプロセッサのデータ転送処理時間が必要になることである。この問題により、ディスク制御装置にかかる負荷は大きく、ディスク制御装置の処理性能の低下につながる。

【0009】この問題を解決する一つの例と考えられる特開平7-281959号公報の方法では、二台の磁気ディスク装置間で互いに各々の動作状態を連絡する手段を磁気ディスク装置内に設け、この連絡手段を用いて、二台の磁気ディスク装置で同時に書き込み動作を進めるという方法を提供している。この方法を用いた場合、磁気ディスク装置とディスク制御装置との間のデータ転送量は一台の磁気ディスク装置に対するデータ転送量と同じである。

【0010】

【発明が解決しようとする課題】従来の技術として説明してきたデータ二重化の方法では、二台の磁気ディスク装置にデータ転送を行うためにディスク制御装置のデータ転送処理負荷が増加する。また、この問題を解決する一例として挙げた特開平7-281959号公報の方法も、磁気ディスク装置に専用の連絡手段と連絡線を設けなければならないことや、データの書き込み時には二台の磁気ディスク装置を占有してしまうことから、多数の磁気ディスク装置を接続して、データ転送の多重度を上げて性能を向上させるディスクアレイ装置においては、磁気ディスク装置の利用効率が低下してしまう。

【0011】本発明の目的は、RAID1の冗長構成を用いるディスクアレイ装置において、データ二重化におけるディスク制御装置と磁気ディスク装置との間のデータ転送量を低減する方法を提供すること、また既存の磁気ディスク装置に対して、この方法のために専用の制御手段などのハードウェアの追加を行わずに、この方法を実現することである。

【0012】

【課題を解決するための手段】上記目的を達成するために、複数のディスク装置と、ホストコンピュータからの入出力要求をうけて該ディスク装置に対するデータ入出

力を制御するディスク制御装置と、該複数のディスク装置と該ディスク制御装置を接続するインターフェースからなり、該ディスク装置の内の複数台が一組となりデータを重複して格納する冗長構成を用いるディスクアレイ装置において、データの書き込み時に、該ディスク制御装置から第一のディスク装置に該インターフェースを介してデータを書き込む手段と、該第一のディスク装置が、該第一のディスク装置が接続しているインターフェース上に存在する同一の冗長構成グループの第二のディスク装置に該データを転送する手段を設ける。

【0013】また、該ディスク装置に、複数のコマンド及びデータを一時的に保持して置く一時記憶手段と、他のコマンド実行中に、同一の冗長構成グループのディスク装置からのコマンドおよびデータを該一時記憶手段にて受信および保持する手段と、該一時記憶手段に保持されている実行待ちの複数のコマンドの実行順序を変更する手段を設ける。

【0014】

【発明の実施の形態】本発明の提供する、RAID1の冗長構成を用いるディスクアレイ装置における、データ二重化の方法を、図1から図7を参照して詳細に説明する。

【0015】図1は、本発明に係わるディスクアレイ装置の一実施例を簡略に示したものである。ここに示した実施例では、ディスク制御装置と磁気ディスク装置とを接続するインターフェースとして、FC-AL (Fibre Channel Arbitrated Loop)を用いた場合について示すが、磁気ディスク装置以外の記憶装置（光磁気ディスク装置等）を接続すること、あるいは、インターフェースとしてFC-AL以外のものを用いて接続しても本発明を実施可能である。

【0016】図1に示すように、ディスクアレイ装置はディスク制御装置101と、二台一組でRAID1冗長構成グループを形成する磁気ディスク装置102a、102bと磁気ディスク装置103a、103bと、該ディスク制御装置と該磁気ディスク装置とを接続するFC-ALインターフェース104から構成される。ディスク制御装置101はホストコンピュータとのインターフェースを制御するインターフェース制御回路105、マイクロプロセッサ106、内部バス107、ホストコンピュータと磁気ディスク装置間で転送されるデータを一時的に保持するキャッシュメモリ108、FC-ALインターフェースの制御を行うFC-AL制御回路109より構成される。マイクロプロセッサ106は図1には示さないメモリ内のプログラムにより動作する。

【0017】ホストコンピュータから書き込み要求が発せられると、まず書き込みデータはキャッシュメモリ108に書き込まれる。マイクロプロセッサ106は、ホストコンピュータの書き込み要求のパラメータに従って、キャッシュメモリ108に保持されている該データをFC

—AL制御回路109に渡すと共に、FC—AL制御回路109に対して、該渡したデータを磁気ディスク装置へ書き込むことを指令する。この時、マイクロプロセッサ106は、書き込み命令を発行した先の磁気ディスク装置（ここでは仮に磁気ディスク装置102aとする）、およびそれと組になっている磁気ディスク装置102b（データ二重化用のディスク）を他のディスク装置から識別するためのID番号と、該データの書込まれるべき論理ブロックアドレス等の制御情報とを同時にFC—AL制御回路109に転送する。この書き込み命令を受けたFC—AL制御回路109は、上記制御情報を基にDDWRITEコマンド（後述する）を作成し、このDDWRITEコマンドをFC—ALインターフェースを介して磁気ディスク装置へ発行することにより、ホストからのデータを磁気ディスク装置へ転送する。

【0018】書き込み命令を受けた磁気ディスク装置102aは転送されてきたデータをディスクに書き込むと共に、該データを組になっている第二の磁気ディスク装置102bに対して転送する。

【0019】図1の実施例では、従来例である図8と比較して、磁気ディスク装置102a、102b、103a、103bが同一のFC—ALインターフェース104上に接続されていることから、本発明の提案する磁気ディスク装置間で二重化のためのデータ転送がディスク制御装置101を経由せずに行う事が可能であり、データ二重化におけるディスク制御装置101と磁気ディスク装置間のデータ転送量の低減が可能となる。

【0020】なお、図1のディスク制御装置と磁気ディスク装置間を接続するインターフェースには、バスのアービトレーション（調停）を行う機能が搭載される。装置としては、ハードウェアとして搭載されやすいものを用い、かつ単位時間におけるデータ転送の多重度が高いものを用いるのが望ましい。このバスのアービトレーション機能は、ディスク制御装置がコマンドを磁気ディスク装置に送信する際や、磁気ディスク装置が二重化データを組になっている二重化ディスク装置へコマンドを送信する際に、使用される。

【0021】図2は、図1の磁気ディスク装置102aの構成の一例を示したものである。ただし、この図1の磁気ディスク装置102a、102b、103a、103bはすべて同一の構成であるとする。磁気ディスク装置102aは、磁気ディスクを制御する磁気ディスク制御部201、データを格納するディスク202、ディスク202とのデータ転送を行う磁気ヘッド203を備えるアーム204からなる。また、磁気ディスク制御部201は、ディスク制御装置101とのインターフェースであるFC—ALの制御を行うFC—AL制御回路213、磁気ディスクの制御を行うマイクロプロセッサ205、アーム204やディスク202の制御を行うディスクコントローラ206、書き込み・読み込みの制御を行う

読み／書き制御回路207、受け付けたコマンドやデータを一時的に保持して置く手段としてのバッファメモリ208からなる。

【0022】図2では、バッファメモリ208には、読み出しデータ用209、書き込みデータ用210のバッファメモリ、および二重化用データ転送先アドレス用のバッファメモリを備える。二重化用データ転送先アドレス用のバッファメモリには、磁気ディスク装置がDDWRITEコマンドを受けたとき、DDWRITEコマンド内に示される二重化用データ転送先アドレスである二重化用磁気ディスク装置のID番号が格納される。この転送先アドレスは、二重化先の磁気ディスク装置へデータを転送する際に使用される。

【0023】また、バッファメモリ208の容量を増やすことによって、バッファメモリ内にインターフェースからの複数のコマンド及びデータを一時的に保持することができる。このため、DDWRITEコマンド以外のコマンド実行中にも、同一の元長構成グループの磁気ディスク装置からのコマンドおよびデータをバッファメモリにて受信および保持することが可能となる。この結果、該バッファメモリに保持されている実行待ちの複数のコマンドの実行順序を変更することができ、必要に応じて二重化データの書込処理を先に実行するなど、また、逆に二重化データの書込処理を後にするなど、通常の処理の発生具合を考慮した二重化データの書込処理が可能となり、柔軟なディスクアレイ装置を構成することができる。

【0024】次に、図3および図4を用いてコマンドの説明をする。

【0025】図3は、本実施例のFC—ALインターフェースにおける、新設のコマンド（DDWRITEコマンドと呼ぶ）の構造を示す。このコマンドは、ディスク制御装置101から磁気ディスク装置102aに対して、データの書き込みを指示し、かつ磁気ディスク装置102aに対して、データの二重化先である磁気ディスク装置102bに該データを転送することを指示する内容を持ち、EC—AL制御回路109によって作成され、発行される。FC—ALインターフェースはSCSI—3規格をサポートしており、これに規定されている、グループ6のベンダが独自に定義してよいオペレーションコードC0H（バイト0）を使用する。このコマンドには、二重化先である磁気ディスク装置102bを示す「転送先磁気ディスク装置SISI ID番号」（図3の8バイト目）を有する。磁気ディスク装置102aは、これを参照することによって、データの二重化先である磁気ディスク装置102bにディスク制御装置101から受けた該データを転送する。

【0026】図4は、既存のコマンドを用いてDDWRITEコマンドの機能を実現したい場合に使用するXORコマンドの構成を示す。XORコマンド（オペレーシ

ョンコード80H)はSCSI-3で規定されており、このコマンド中の未使用ビット(図4のバイト1の5及び6ビット目)を排他的論理和演算の実行の有無に対応させ、排他的論理和演算の実行の無しを設定することで、DDWRITEコマンドの機能相当を実現可能となる。

【0027】排他的論理和(XOR:exclusive OR)は、先述のディスクアレイ装置にいくつかある冗長構成の内、RAID4とRAID5と呼ばれるものに必要なパリティデータ(故障時のデータ回復用の冗長データ)を生成する為に必要な演算である。データをブロック単位で複数の磁気ディスク装置に格納するRAID4、RAID5では、冗長構成をとる複数の磁気ディスク装置内のデータをブロック単位で排他的論理和を演算し、この演算結果をパリティデータとして磁気ディスク装置に格納する。

【0028】図6は、本来のXORコマンドによるディスクアレイ装置の動作とコマンドやデータの転送経路とを簡略に示した図である。図6中では磁気ディスク装置602a、602b、602c、602dの四台の磁気ディスク装置でRAID5の冗長構成をなす1つのグループを形成しているとする。以下、XORコマンドによるディスクアレイ装置の動作を簡単に説明する。

【0029】ホストコンピュータのデータ書き込み要求を受けたディスク制御装置601は、XORコマンドを該データを書き込むべき磁気ディスク装置602aに発行する。磁気ディスク装置602aは書き換えられる旧データを格納している磁気ディスク装置であり、一時保持手段であるバッファメモリに新データが書き込まれる(経路604)と同時に、磁気ディスク装置602aは旧データをディスクから読み出し、該新データと該旧データの排他的論理和演算を磁気ディスク装置602a内で行い、該新データをディスクに書き込んだ後、該新データと該旧データの排他的論理和の結果である中間データを、旧パリティデータを格納している磁気ディスク装置602dに転送する(経路605)。

【0030】該中間データは該旧パリティデータを格納している磁気ディスク装置602dのバッファメモリに格納される。その後は、先述の動作と同様に、中間データと旧パリティデータの排他的論理和が演算され、該演算の結果である新パリティデータがディスクに書き込まれ、新パリティデータの書き込みの正常終了ステータスをパリティデータを格納する磁気ディスク装置602dからデータを格納する磁気ディスク装置602aに返送し、該正常終了ステータスを受けた磁気ディスク装置602aは、XORコマンドが正常に終了したことをディスク制御装置601に返送する。

【0031】ここで、XORコマンドによる動作の内、パリティデータ生成のための旧データおよび旧パリティデータのディスクからの読み出し、および排他的論理和

演算を実行させなければ、旧データが格納されているディスク装置602aでは、ディスク制御装置601から転送された新データが、当該ディスク制御装置602aに書き込まれると共に、そのまま中間データとして旧パリティデータが格納されている磁気ディスク装置602dへ転送されることになる。そして磁気ディスク装置602dでは、中間データがそのまま磁気ディスク装置602dに書き込まれることになる。

【0032】即ち、図4に示すXORコマンドにおいて、1バイト目の第6、5ビットが未使用となっているので、ここを排他的論理和演算の実行の有無に対応させ、排他的論理和演算の実行の無に設定し、当該XORコマンドを実行させれば、DDWRITEコマンドと同様に、ディスク制御装置から転送されたデータが、旧データの格納されている磁気ディスク装置602aとパリティデータの格納されている磁気ディスク装置602dに転送されることになる。

【0033】図5は、DDWRITEコマンドによる、ディスクアレイ装置の動作とコマンドやデータの転送経路とを示した図である。経路501~505の順でコマンド又はデータが転送される。また、図7はDDWRITEコマンドを受けた磁気ディスク装置のシーケンスを示したフローチャートである。以下、図5および図7を用いて、DDWRITEコマンドによる動作を説明する。

【0034】ホストコンピュータのデータ書き込み要求を受けたディスク制御装置101は、DDWRITEコマンドを該データを書き込むべき磁気ディスク装置102aに発行する。ここで、ディスク制御装置101は、DDWRITEコマンドを発行する磁気ディスク装置を二台一組になっている磁気ディスク装置のどちらにしてもよい。常にどちらか一方に発行するように決めておいてもよいし、ランダムに二台のどちらかを選ぶ方法でもよい。望ましい実施例としては、磁気ディスク装置に対して、既に発行されて実行を待っているコマンドの待ち行列をディスク制御装置101が監視していて、待ち行列の短い方にコマンドを発行する方法である。

【0035】また、すでに待ち行列に入っている実行待ちコマンド列の実行順序を並べ変える機能はSCSI仕様に規定されているキューイング機能を用いて、コマンドのパラメータの示す論理ブロックアドレスと転送ブロック数により、シーク・サーチ動作が最小となるように最適化する手法を用いて実現すれば良い。

【0036】コマンドを受けた磁気ディスク装置102aは、シーク・サーチ動作を開始する(ステップ701、702)。同時に、ディスク制御装置から転送される新データを一時保持手段であるバッファメモリに格納し(ステップ703)、データ書き込み可能な状態になるまで待つ(ステップ705)。シーク・サーチ動作が完了し(ステップ704)、ディスクへデータ転送可能状

態になると、磁気ディスク装置はバッファメモリより該新データをディスクに書き込む（ステップ706）。次に、データの二重化のために、磁気ディスク装置102aは、同一のRAID1グループを構成するもう一方の磁気ディスク装置102bに該新データを転送する。この時に、磁気ディスク装置102aが磁気ディスク装置102bへ発行するコマンドは、ディスク制御装置が磁気ディスク装置に書込を指示するところの通常の手書き込みコマンドと同様でよい（ステップ707）。該新データをディスクに書き込むまでの動作は最初にコマンドを受けた磁気ディスク装置と同様である（ステップ708～713）。

【0037】DDWRITEコマンドによる該新データのディスクへの書き込みが完了した場合、書き込みの正常終了ステータスをデータを格納する磁気ディスク装置102aからデータを格納する磁気ディスク装置102bに転送し、該正常終了ステータスを受けた該磁気ディスク装置102bは、DDWRITEコマンドが正常に終了したことをディスク制御装置101に返送する。この場合、ディスク制御装置101は、一度正常終了を示すステータスを受けるだけで二重データ書込が終了したことを知ることができ、個々の磁気ディスク装置から終了ステータスを受けていた場合と比べて、ステータスを受ける回数が少なくなり、その管理が容易となり、また、オーバーヘッドを削減できる。あるいは、二台の磁気ディスク装置102a、102bが、それぞれディスクへの書き込み動作が正常に終了した時点で、正常終了ステータスをディスク制御装置101へ直接返送してもよい。この場合、ディスク制御装置101がDDWRITEコマンドを正常終了と判断するのは、二台の磁気ディスク装置102aおよび102bから正常終了ステータスを受けた時点となる。

【0038】

【発明の効果】本発明により、RAID1の冗長構成を用いるディスクアレイ装置において、ディスク制御装置と磁気ディスク装置間のデータ転送量を低減させ、処理性能を向上させる効果がある。また、少数のインターフェースで本発明は効果を発揮し、本発明のために特に特別な制御手段などを必要としないため安価に高性能なディスクアレイを実現できる。

【図面の簡単な説明】

【図1】本発明に関するディスクアレイ装置の一構成例を示した概略図。

【図2】本発明に関する磁気ディスク装置の一構成例を示した概略図。

【図3】本発明に関する磁気ディスク装置間でデータの二重化を実行させるコマンドの一構成例を示す構成図。

【図4】本発明に係わり、RAID4又は5の冗長構成のディスクアレイ装置に用いられるXORコマンドの構成図。

【図5】本発明に係わるディスクアレイ装置の動作の一例を示す概略図。

【図6】本発明に関するディスクアレイ装置の動作の一例を示す概略図。

【図7】本発明に関するディスクアレイ装置の動作の一例を示すフローチャート。

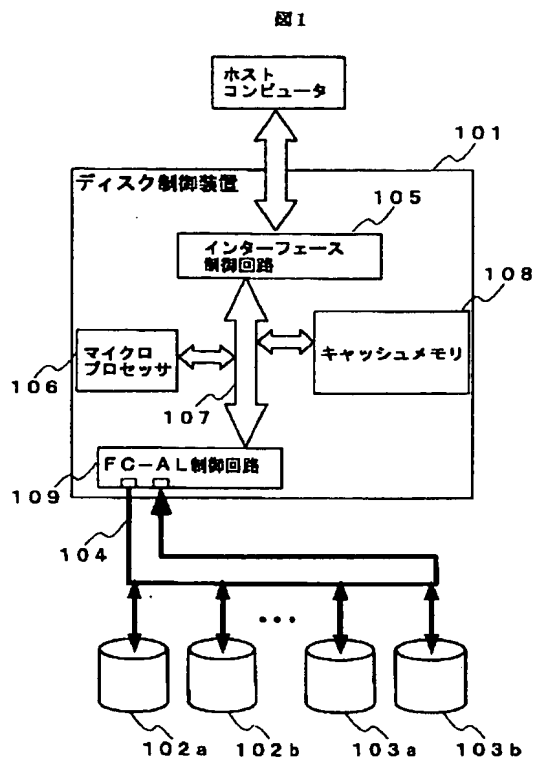
【図8】従来のディスクアレイ装置の構成を示した概略図。

【図9】従来のディスクアレイ装置の動作を示すフローチャート。

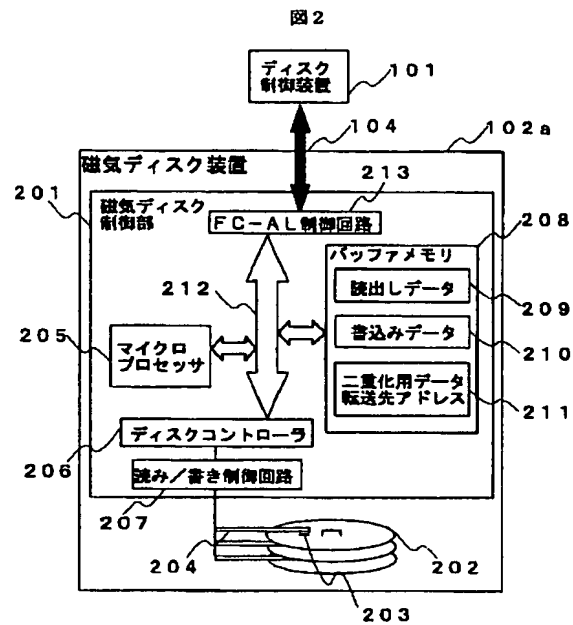
【符号の説明】

101、601、801…ディスク制御装置、102a、102b、103a、103b、602a、602b、602c、602d、802a、802b、803a、803b…磁気ディスク装置、804a、804b、804c、804d…SCSIインターフェース、104、603…FC-ALインターフェース、105、805…インターフェース制御回路、106、205、806…マイクロプロセッサ、107、212、807…内部バス、108、808…キャッシュメモリ、109、213…FC-AL制御回路、201…磁気ディスク制御部、202…ディスク、203…磁気ヘッド、204…アーム、206…ディスクコントローラ、207…読み／書き制御回路、208…バッファメモリ、209…読出しデータ用バッファメモリ、210…書き込みデータ用バッファメモリ、211…二重化用データ転送先アドレス用バッファメモリ、301…SCSIコマンド、401…XORコマンド、809a、809b、809c、809d…SCSI制御回路。

【図1】



【図2】

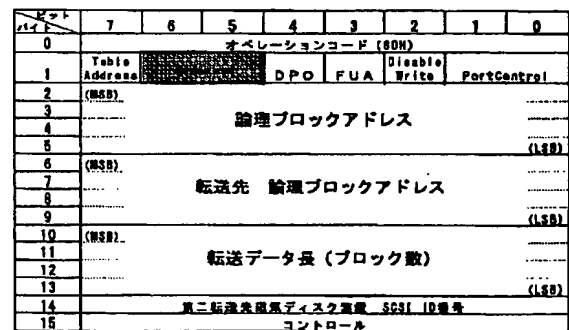
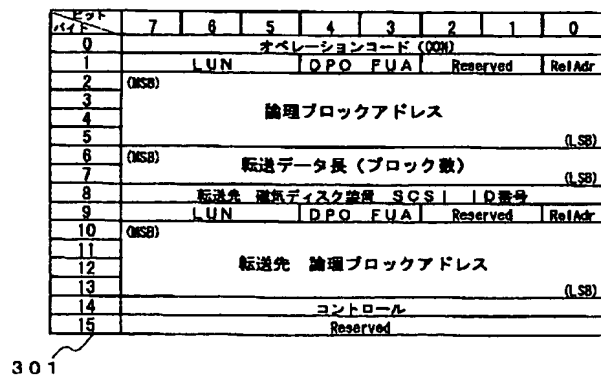


【図4】

図4

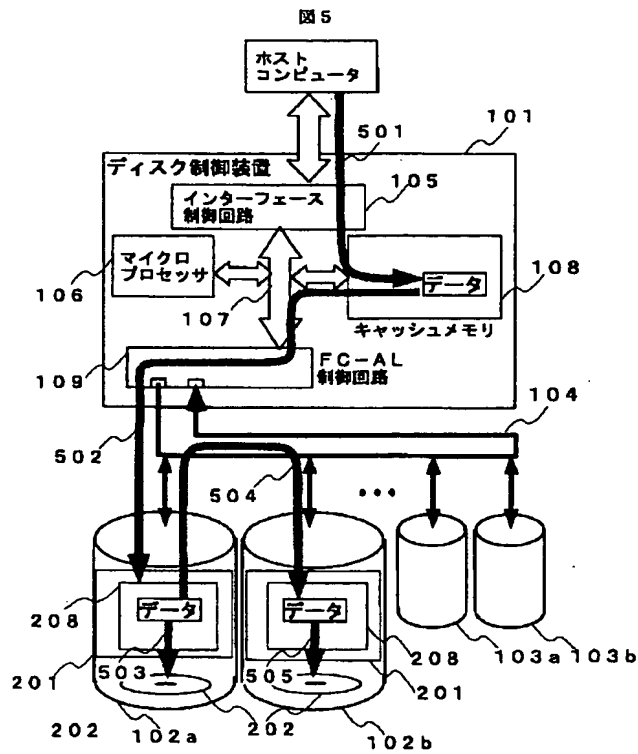
【図3】

図3

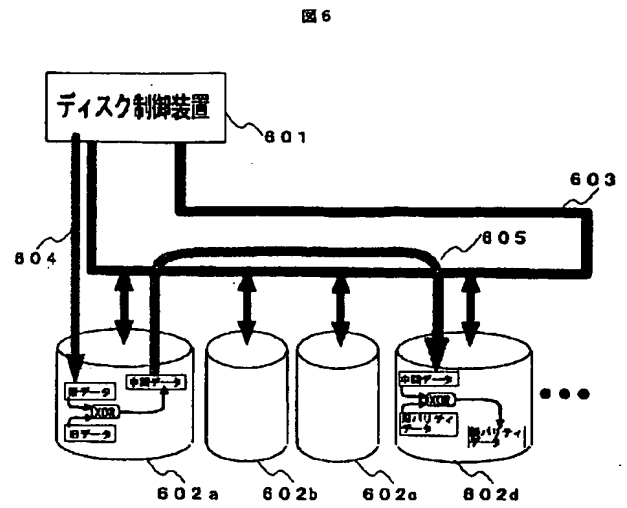


401

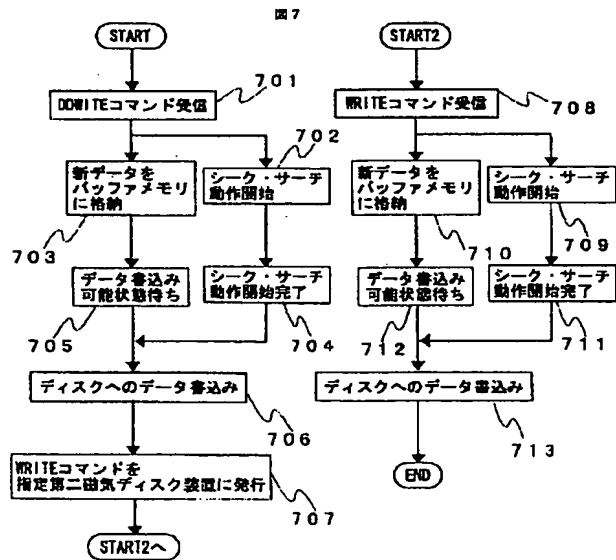
【図5】



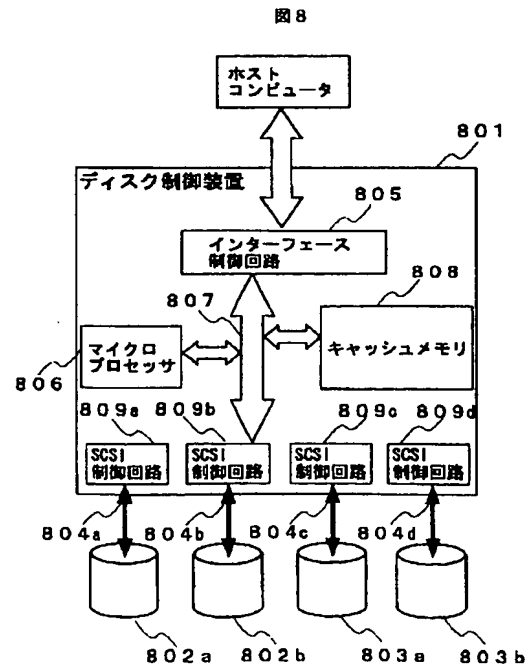
【図6】



【図7】

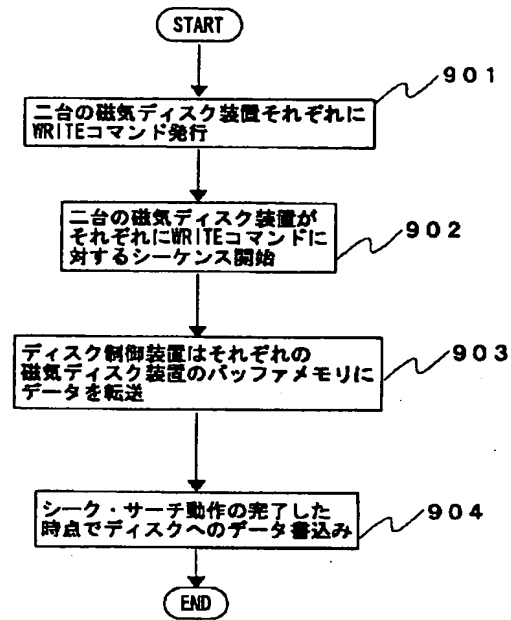


【図8】



【図9】

図9



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.